Journal Name

COMMUNICATION

Aqueous Recognition of Purine and Pyrimidine Bases by an Anthracene–Based Macrocyclic Receptor

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

www.rsc.org/

Danny Van Eker, Soumen K. Samanta, Anthony P. Davis,*

A water-soluble bis-anthracenyl tetralactam binds biogenic heterocycles with high affinities in aqueous solution, rising to 10^7 M^{-1} for the purine hypoxanthine. Recognition occurs through a combination of hydrogen bonding and hydrophobic interactions, and results in fluorescence changes which suggest applications in sensing.

Controlling molecular recognition in water is a key objective of supramolecular chemistry.[1] Water is the solvent of life, so by far the most relevant for biological and medical applications. Moreover its effects are complex, depressing polar interactions (e.g. hydrogen bonding) through solvation of binding groups yet driving association through hydrophobic interactions between apolar surfaces. The recognition of polar organic molecules in water is especially challenging, as both factors come into play. From biology we can learn that high affinities are possible if both H-bonding and hydrophobic moieties are properly deployed (sometimes very high, as in the case of the avidin-biotin pairing[2]). However, there are still relatively few synthetic receptors where both types of interaction are preorganised to achieve molecular recognition in water.[3]

Among substrates of biological interest, there are many that include purine or pyrimidine base units. These structures have been addressed previously with some success, mostly through the application of hydrophobic interactions supplemented by electrostatic attraction in the case of charged substrates (ATP, nucleic acids etc.).[4,5] Hydrogen bonding has not generally been exploited, presumably because it is considered ineffective in aqueous media. For uncharged substrates, affinities have generally been in the range 10^3 - 10^5 M^{-1}.[5] We now report that the bis-anthracenyl tetralactam **1** (Fig. 1), with inward-directed NH groups, serves as a receptor for several of these heterocycles in water. The receptor is notably selective, favouring the important biomarker hypoxanthine **3** which is bound with K_a ~10^7 M^{-1}. ^{1}H NMR and modelling studies suggest that hydrogen bonding contributes to this unusually high affinity, illustrating the potential for combining polar and apolar interactions to achieve molecular recognition in water.

Receptor **1** belongs to a family of tetralactams, with identical core structures but different solubilising groups, which were originally designed to bind glucose in sandwich-type complexes.[6,7] Glucose was indeed bound with K_a = 60-90 M^{-1} in water, and the complex geometry was confirmed using X-ray crystallography.[8] Moreover the anthracene fluorescence was enhanced on binding, suggesting the potential for glucose sensing. Unfortunately, the sensitivity to glucose was lost when the system was transferred to blood serum, and this raised the possibility that other biogenic small molecules might bind to the cavity.

Considering the likely interferents that could be present, polar aromatic compounds seemed good candidates. Modelling suggests that the anthracene planes in **1** are ~7.3 Å apart. This spacing is approximately twice the inter-base distance in DNA, so almost ideal for an aromatic guest[9] (and slightly too small for glucose[10]). Purines and pyrimidines possess lone pairs in the plane of the aromatic ring, with potential for H-bonding to receptor NH, so could be good substrates for **1**.[11]

Uric acid **2** is present in serum at relatively high levels (typically ~0.3 mM) and seemed a likely possibility. Indeed, both ^{1}H NMR and fluorescence titrations gave evidence that **1** binds **2** quite strongly. Titration of uric acid into **1** in D_2O (10 mM phosphate buffer, pH 7.4) caused broadening of the receptor ^{1}H NMR signals and the appearance of new peaks, saturating at ~1.6 equivalents (Fig. S20). The fluorescence emission of **1** in H_2O (buffered similarly) was reduced by addition of **2**, as shown in Fig. 2. The NMR data could not be quantified due to signal broadening, but the fluorescence output fit well to a 1:1 binding model with $K_a = 1.7 \times 10^5$ M^{-1} (Fig 2). Isothermal titration calorimetry (ITC) yielded a similar value of 1.8×10^5 M^{-1}.

School of Chemistry, University of Bristol,
Cantock's Close, BS8 1TS, Bristol, UK
E-mail: Anthony.Davis@bristol.ac.uk

†Electronic Supplementary Information (ESI) available: Details of fluorescence, NMR and ITC measurements, spectra and binding analysis curves. See DOI: 10.1039/x0xx00000x

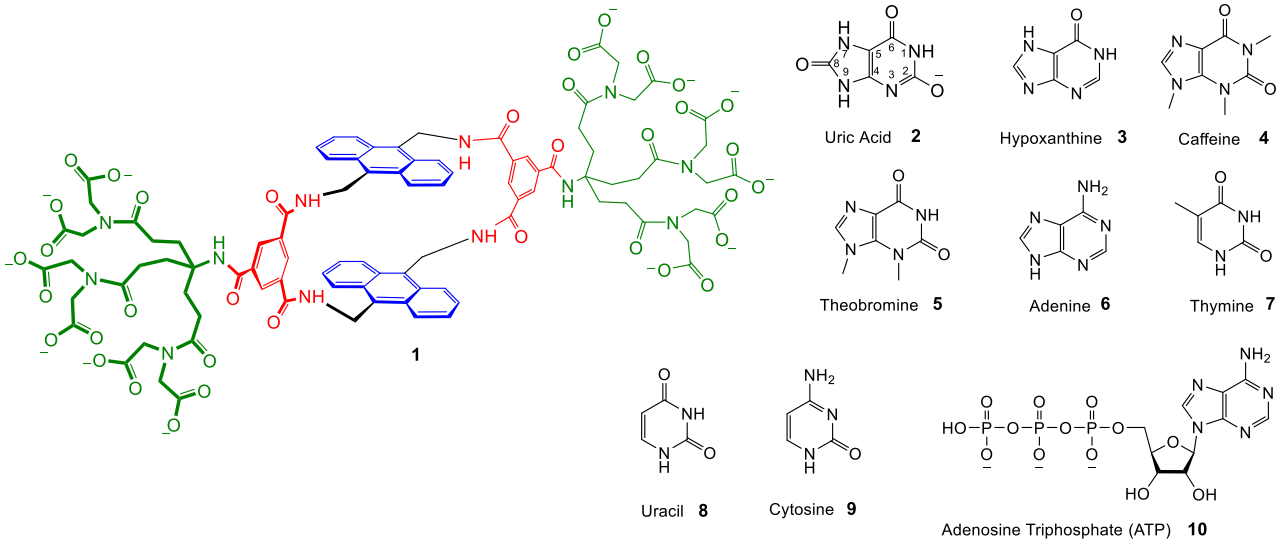

Fig. 1 Receptor **1** and the purine and pyrimidine substrates used in this work. Uric acid **2** is represented as the monoanion.

The above results confirmed that uric acid **2** was a good candidate for the interfering species in blood serum, but also suggested that other purines and/or pyrimidines could be substrates for **1**. We therefore tested the binding of **1** to eight further biogenic heterocycles (**3**-**10**; see Fig. 1). The results are summarised in Table 1. Trianionic ATP **10** was only bound very weakly, but all the neutral substrates were bound with $K_a = 2 \times 10^3$ M^{-1} or greater. Most notably, the titration of hypoxanthine **3** into receptor **1** caused an increase in fluorescence which was analysed to give $K_a = 8.7 \times 10^6$ M^{-1}, ≥ 15 times higher than for other substrates (Fig. 3). Again this result was supported by ITC, which gave $K_a = 8.1 \times 10^6$ M^{-1}. ^{1}H NMR titrations showed changes consistent with binding (Fig S22; see discussion below) but quantification was again prevented by signal broadening. Hypoxanthine is the metabolic product of ATP under oxygen-deficiency, and an important biomarker for conditions like acute cardiac ischemia, prostate cancer and neonatal hypoxia.[12] It is also of interest as an indicator of ageing in meat and fish.[13]

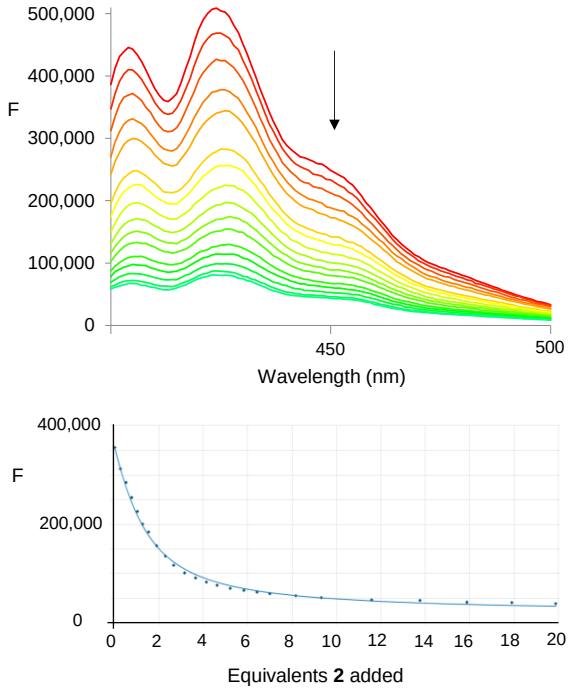

Fig. 2 Top: Fluorescence titration of receptor **1** (11.8 μM) with uric acid **2** in 10 mM phosphate buffer, pH 7.4. F = fluorescence emission intensity. Bottom: Plot of experimental data (dots) and calculated values (line) from 1:1 binding analysis using Bindfit software. Emission observed at 427 nm. $K_a = 1.7 \times 10^5$ M^{-1} ± 1.2%

Table 1. Association constants measured for receptor **1** to heterocyclic guests in aqueous solution.[a]

Guests	K_a [M^{-1}]
Uric Acid **2**	1.7×10^5, 1.8×10^5 [b]
Hypoxanthine **3**	8.7×10^6, 8.1×10^6 [b]
Caffeine **4**	1.7×10^5, 1.6×10^5 [b]
Theobromine **5**	1.1×10^5
Adenine **6**	1.3×10^5
Thymine **7**	5.8×10^5 [c]
Uracil **8**	9.1×10^3 [c]
Cytosine **9**	2.3×10^3
ATP **10**	24

[a] Measured by fluorescence titration in aqueous phosphate buffer (10 mM, pH 7.4) at 298 K, unless otherwise indicated. All data were fitted to a 1:1 binding model. [b] Measured by ITC in aqueous phosphate buffer (10 mM, pH 7.4) at 298 K. [c] Measured by ^{1}H NMR titration in D_2O (10 mM phosphate buffer, pH 7.4) at 298 K. These substrates are thought to be capable of 1:2 binding, so the listed 1:1 K_a values are considered "apparent".

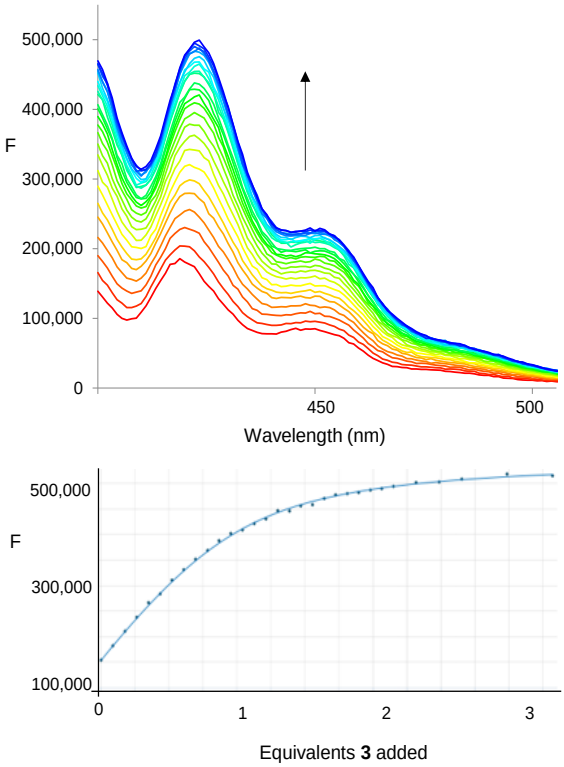

Fig. 3 Top: Fluorescence titration of receptor **1** (550 nM) with hypoxanthine **3** in 10 mM phosphate buffer, pH 7.4. F = fluorescence emission intensity. Bottom: Plot of experimental data (dots) and calculated values (line) from 1:1 binding analysis using Bindfit software. Emission observed at 427 nm. K_a = 8.7×10^6 M^{-1} ± 2.5%.

Of the other substrates, the purine alkaloids caffeine **4** and theobromine **5** also increased fluorescence emission, consistent with K_a = 1.7 × 10^5 M^{-1} and 1.1 × 10^5 M^{-1} respectively. Interestingly, theobromine **5** caused a remarkable 15-fold enhancement (Fig S8), much higher than the ~3-fold increase for caffeine or hypoxanthine. The affinity of **1** for caffeine was confirmed by ITC; a ^{1}H NMR titration again gave broadened peaks. Adenine **6** caused the fluorescence spectrum to shift with a small decrease in intensity, analysed to give K_a = 1.3 × 10^5 M^{-1}. The behaviour of the pyrimidines, thymine and uracil was more complex. In fluorescence titrations the curves showed sigmoidal shapes, consistent with strong 1:1 binding causing minimal changes in emission followed by 1:2 complex formation causing substantial increases (Figs S13 and S15). These substrates were also studied using ^{1}H NMR titrations for which, unlike **2**-**4**, signals remained sharp but moved as expected for binding with fast exchange. Here, the changes fitted well to a 1:1 binding model consistent with K_a = 5.8 × 10^5 M^{-1} for thymine **7** and 9.1 × 10^3 M^{-1} for uracil **8**. Given the evidence for 1:2 binding, which could affect the analyses, these values should be considered "apparent". Finally, cytosine **9** gave small fluorescence changes analysed to give K_a = 2.3 × 10^3 M^{-1}.

The structures of the complexes were studied using ^{1}H NMR and molecular modelling. Of particular interest was the role of hydrogen bonding in promoting binding, especially to hypoxanthine **3**. The NMR titration for **1** + **3** was repeated in H_2O/D_2O 9:1, so that signals due to NH could be followed.[14] Although broadening complicated the interpretation, signals at 8.75 and 9.25 p.p.m. could be assigned to the lactam NH in receptor and complex respectively. The 0.5 p.p.m. downfield shift suggests significant NH···X hydrogen bonding in the complex, given the extensive NH solvation in the free receptor. The movement is twice as large as that observed earlier for **1** + methyl β-D-glucoside in water.[6]

Modelling of the complexes between **1** and purines **2**-**6** was performed using Monte Carlo Molecular Mechanics (MCMM), allowing the guests to move within the cavity before each minimisation. All calculations yielded structures with four intermolecular NH···N or NH···O hydrogen bonds in the range ~2.0-2.2 Å. The ground state structure for the hypoxanthine complex **1**·**3** is shown in Figure 4. Overlap between the host and guest aromatic surfaces is excellent, although this is also true of the urate and adenine complexes **1**·**2** and **1**·**6** (by contrast, caffeine and theobromine protrude slightly from the cavity, see Figs. S35 and S36). The preference for hypoxanthine presumably results from subtle differences in guest geometry and solvation. Of the pyrimidine guests, thymine **7** and uracil **8** were modelled binding to **1** in 2:1 stoichiometry. Both these guests can form H-bonded dimers. Insertion into the cavity of **1** followed by energy minimisation gives plausible structures held together by six intermolecular hydrogen bonds. The complex between **1** and the uracil dimer

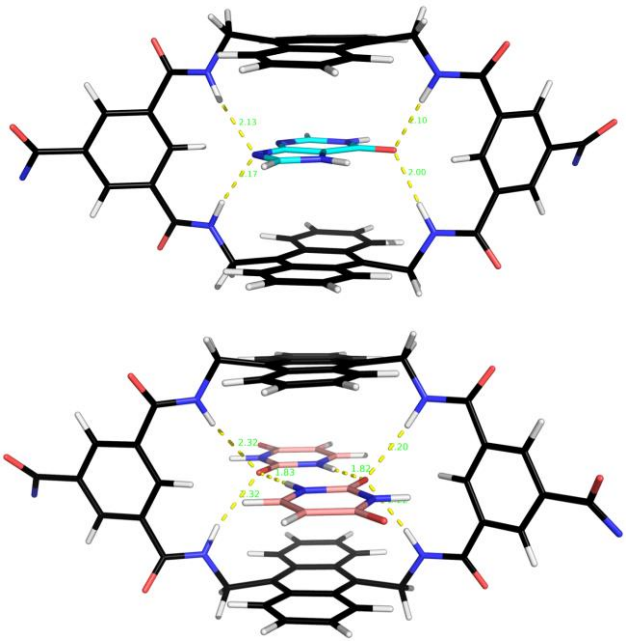

Fig. 4 Top: Model of hypoxanthine **3** bound to receptor **1**. Side-chains are omitted for clarity. Intermolecular hydrogen bonds are shown as yellow dashed lines. Side-chains are omitted for clarity. Bottom: Model of uracil **8** dimer bound to receptor **1**. For details of calculations see supporting information.

is shown in Figure 4. Finally, an MCMM calculation on the cytosine complex **1**·**9** gave a structure with just two intermolecular H-bonds, consistent with the relatively low binding constant.

In conclusion we have shown that while macrocycle **1** is a moderate receptor for glucose, it is highly effective for several biogenic heterocycles. Affinities rise to 10^7 M^{-1} for the optimal substrate hypoxanthine **3**, unusually high for a synthetic receptor binding a biologically relevant, neutral polar molecule in water. While this discovery limits the potential of **1** as a glucose sensor, the selectivity for **3** and fluorescence response raises the possibility of hypoxanthine sensing, with applications in medicine and food safety. Meanwhile, more generally, the results provide further evidence that hydrogen bonding and hydrophobic interactions can combine to achieve strong and selective binding in water.

We are grateful to the EPSRC-funded Centre for Doctoral Training in Chemical Synthesis (EP/L015366/1) and Carbometrics Ltd. for a studentship to DVE.

Conflicts of interest

There are no conflicts to declare.

Notes and references

1. (a) *Supramolecular Chemistry in Water*, ed. S. Kubik, Wiley-VCH, Weinheim, 2019. (b) P. S. Cremer, A. H. Flood, B. C. Gibb and D. L. Mobley, *Nature Chem.*, 2018, **10**, 8-16. (c) A. P. Davis, S. Kubik and A. D. Cort, *Org. Biomol. Chem.*, 2015, **13**, 2499-2500. (d) E. A. Kataev and C. Muller, *Tetrahedron*, 2014, **70**, 137-167. (e) G. V. Oshovsky, D. N. Reinhoudt and W. Verboom, *Angew. Chem., Int. Ed.*, 2007, **46**, 2366-2393.
2. W. Q. Liu, S. K. Samanta, B. D. Smith and L. Isaacs, *Chem. Soc. Rev.*, 2017, **46**, 2391-2403.
3. Selected references: (a) C. Allott, H. Adams, P. L. Bernad, C. A. Hunter, C. Rotger and J. A. Thomas, *Chem. Commun.*, 1998, 2449-2450. (b) M. Fokkens, T. Schrader and F. G. Klarner, *J. Am. Chem. Soc.*, 2005, **127**, 14415-14421. (c) C. Schmuck, D. Rupprecht and W. Wienand, *Chem. Eur. J.*, 2006, **12**, 9186-9195. (d) B. Verdejo, G. Gil-Ramirez and P. Ballester, *J. Am. Chem. Soc.*, 2009, **131**, 3178-3179. (e) E. M. Peck, W. Liu, G. T. Spence, S. K. Shaw, A. P. Davis, H. Destecroix and B. D. Smith, *J. Am. Chem. Soc.*, 2015, **137**, 8668-8671. (f) T. J. Mooibroek, J. M. Casas-Solvas, R. L. Harniman, C. M. Renney, T. S. Carter, M. P. Crump and A. P. Davis, *Nature Chem.*, 2016, **8**, 69-74. (g) A. Lascaux, G. De Leener, L. Fusaro, F. Topic, K. Rissanen, M. Luhmer and I. Jabin, *Org. Biomol. Chem.*, 2016, **14**, 738-746. (h) G. B. Huang, S. H. Wang, H. Ke, L. P. Yang and W. Jiang, *J. Am. Chem. Soc.*, 2016, **138**, 14550-14553. (i) H. Yao, H. Ke, X. B. Zhang, S. J. Pan, M. S. Li, L. P. Yang, G. Schreckenbach and W. Jiang, *J. Am. Chem. Soc.*, 2018, **140**, 13466-13477. (j) W. Q. Liu, A. Johnson and B. D. Smith, *J. Am. Chem. Soc.*, 2018, **140**, 3361-3370. (k) O. Francesconi, M. Martinucci, L. Badii, C. Nativi and S. Roelens, *Chem. Eur. J.*, 2018, **24**, 6828-6836. (l) R. A. Tromans, T. S. Carter, L. Chabanne, M. P. Crump, H. Li, J. V. Matlock, M. G. Orchard and A. P. Davis, *Nature Chem.*, 2019, **11**, 52-56. (m) G. Peñuelas-Haro and P. Ballester, *Chem. Sci.*, 2019, **10**, 2413-2423. (n) A. P. Davis, *Chem. Soc. Rev.*, 2020, **49**, 2531-2545.
4. A. M. Agafontsev, A. Ravi, T. A. Shumilova, A. S. Oshchepkov and E. A. Kataev, *Chem. Eur. J.*, 2019, **25**, 2684-2694. I. Pont, C. Galiana-Rosello, A. Lopera, J. González-García and E. García-España, in ref. 1a. p. 115-160. J. Oh, H.-W. Rhee and J.-I. Hong, in *Synthetic Receptors for Biomolecules : Design Principles and Applications*, ed. B. D. Smith, Royal Society of Chemistry, Cambridge, 2015, p. 204-252.
5. J. Lagona, B. D. Wagner and L. Isaacs, *J. Org. Chem.*, 2006, **71**, 1181-1190. H. Abe, Y. Mawatari, H. Teraoka, K. Fujimoto and M. Inouye, *J. Org. Chem.*, 2004, **69**, 495-504. P. Cudic, M. Zinic, V. Tomisic, V. Simeon, J.-P. Vigneron and J.-M. Lehn, *J. Chem. Soc., Chem. Commun.*, 1995, 1073-1075. S. Claude, J.-M. Lehn, F. Schmidt and J.-P. Vigneron, *J. Chem. Soc., Chem. Commun.*, 1991, 1182-1185.
6. C. Ke, H. Destecroix, M. P. Crump and A. P. Davis, *Nature Chem.*, 2012, **4**, 718-723.
7. H. Destecroix, C. M. Renney, T. J. Mooibroek, T. S. Carter, P. F. N. Stewart, M. P. Crump and A. P. Davis, *Angew. Chem., Int. Ed.*, 2015, **54**, 2057-2061.
8. P. K. Mandal, B. Kauffmann, H. Destecroix, Y. Ferrand, A. P. Davis and I. Huc, *Chem. Commun.*, 2016, **52**, 9355-9358.
9. Organic-soluble relatives of **1** were known to accept aromatic and other planar substrates. See: D. H. Li and B. D. Smith, *Beilstein J. Org. Chem.*, 2019, **15**, 1086-1095.
10. R. A. Tromans, T. S. Carter, L. Chabanne, M. P. Crump, H. Li, J. V. Matlock, M. G. Orchard and A. P. Davis, *Nature Chem.*, 2019, **11**, 52-56.
11. In separate work, it was also found that bis-anthracenyl tetralactams such as **1** bind synthetic squaraine dyes very strongly in water. See for example refs. 3e and 3j.
12. O. D. Saugstad, *Pediatric Research*, 1988, **23**, 143-150.
13. H. S. Nakatani, L. V. dos Santos, C. P. Pelegrine, S. T. Marques Gomes, M. Matsushita, N. E. de Souza and J. V. Visentainer, *Am. J. Biochem. Biotechnol.*, 2005, **1**, 85-89. J. A. V. Albelda, A. Uzunoglu, G. N. C. Santos and L. A. Stanciu, *Biosens. Bioelectron.*, 2017, **89**, 518-524.
14. Similar titrations were performed for **2** and **4**, but in these cases the lactam NH signals were either lost though broadening or could not be identified in the complex.

